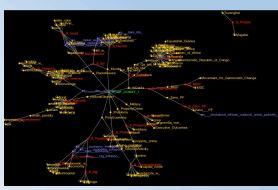
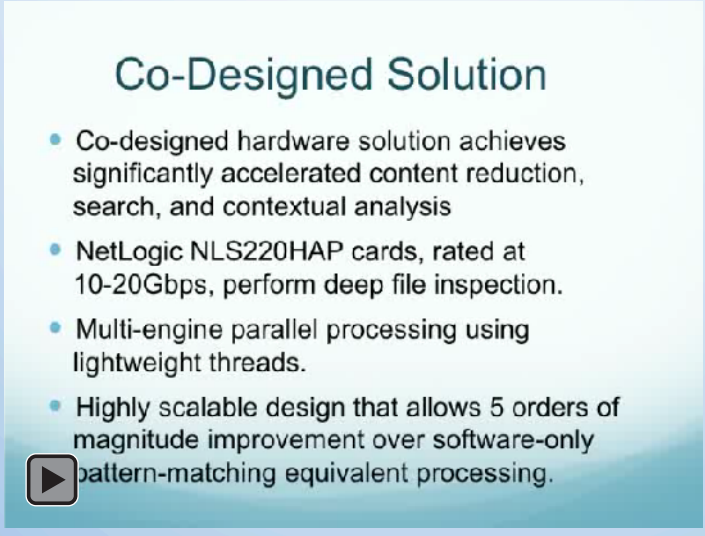
Indago: Digital Knowledge Discovery

The Power of Contextual Analysis



20:1 reduction of content allows analysis of millions of entries (e.g., cataloging of bibliographic records, email traffic analysis, etc.) in real time



- The most common and simplest form of analyzing a large repository uses simple word-occurrence searching techniques in which a set of terms are used to search for relevant content
- A significant drawback of this approach is that the results contain terms in the wrong context−for example, the term "Apple" gets results related to fruits and computers
- LANL's Digital Knowledge Discovery (DKD) team is creating complex models of language contained in unstructured text to build increasingly sophisticated tools that move beyond word occurrence to contextual analysis, giving us the ability to discover patterns found in large collections

Jorge Roman, jhr@lanl.gov and phone
Dave DuBois, dhd@lanl.gov and phone

Los Alamos NATIONAL LABORATORY EST. 1943

NNSA National Nuclear Security Administration